

REMARKS

This Amendment responds to the final office action mailed on November 15, 2007. Claims 1, 10, 19, 28 and 35 are amended. Claims 2-8, 11-18, 21-27, 29-34 and 37-39 have been cancelled. Dependent claims 40-42 have been added, to define an additional aspect of applicants' disclosed embodiment. Claims 1, 9, 10, 19, 20, 28, 35, 36 and 40-43 remain pending. Claims 1, 9, 10, 19, 20, 28, 35 and 36 stand rejected. Reconsideration is respectfully requested in light of the above amendments and the following remarks.

Claim Rejections

In the non-final office action, claims 1, 9-10, 19-20, 28, and 35-36 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Reiley et al (U.S. 2002/0016801) in view of Bourbakis (U.S. 2003/0145279). Dependent claim 26 has been rejected under 35 U.S.C. § 103(a) as being unpatentable over Reiley et al and Bourbakis, as applied to claims 1, 9-10, 19-20, 28, and 35-36, and further in view of Whitledge et al (U.S. Patent 6,925,595).

Claim 26 has been cancelled to reduce the issues.

The patent owner respectfully disagrees with the rejection of claims 1, 9-10, 19-20, 28, and 35-36 were rejected under 35 U.S.C. § 103(a). Nonetheless, in an effort to expedite prosecution of the instant application, the claims have been amended to further distinguish over the cited references, as discussed in greater detail below.

With respect to independent claims 1, 10 and 19 of the instant application, the method disclosed in Reiley is specifically limited to parsing a Web document to create a data structure

containing a hierarchical organization of elements from the Web document (claim 1 of Reiley).

The passages from Reiley cited in the office action also acknowledge that Reiley is limited to documents that have such a hierarchical structure. For example, paragraph [0015] states:

According to one aspect of the invention, a content transformer transforms a Web document from a first format into a second format. The content transformer retrieves a copy of the Web document, wherein the Web document comprises one or more elements that are delimited and identified by tags within the Web document; parses the Web document to create a first data structure comprised of a first hierarchical organization of elements from the Web document; conducts a semantic analysis of the elements in the data structure; and re-arranges the elements in the first data structure based upon the semantic analysis to form a second data structure comprised of a new hierarchical organization of elements from the Web page, wherein the new hierarchical organization differs from the first hierarchical organization.

while paragraph [0065] states:

The re-arrangement may include re-organization of the nodes in the hierarchy, removal of one or more nodes from the hierarchy, merging of nodes, and the addition or revision of node identifiers. The semantic analysis and re-arrangement preferably results in a transformed hierarchical structure that properly reflects the hierarchy of the elements of the content. The operations represented by flow diagram boxes 530 and 535 are preferably recursively performed on the hierarchical structure.

As these passages and the claim language make clear, the method in Reiley is limited to analyzing documents with a hierarchical structure. While the method disclosed in the instant application may analyze documents with a hierarchical structure, it also analyzes the content properties and content formatting and can use such data to generate content summaries.

Examiner notes that Reiley discloses determining if the electronic document has a predetermined content structure and in response selecting the plurality of summary entries from the electronic document based on the predetermined **content structure**, citing page 2, paragraph [0017]. The passage cited by the examiner summarizes operation of a content transformer of

Reiley for performing an analysis of the elements of a Web document, the analysis taking into account a structural arrangement of the elements, and rearranging (summarizing) the elements as a result of the analysis to generate a hierarchical data structure that represents the Web document. The content transformer then generates the user device formatted version of the Web document based upon a hierarchical data structure.

Examiner also argues that Reiley discloses selecting the plurality of summary entries from the electronic document based on differences in the **content properties**, citing page 3, paragraphs [0038] – [0041]. Specifically, the examiner argues that he cited passage teaches: “the Web page (electronic document) is divided into several elements including headings, paragraphs, lists, separators, graphics, tables, table items, etc.”. The examiner argues that these are “content properties”, within the meaning of applicants’ claims, and that the transformer “uses analysis rule to categorize the elements...”

The examiner has misquoted the passage from Reiley at page 3, paragraphs [0038] – [0041], which in fact reads: “the Web page 205 is divided into several logical structures or elements, including headings, paragraphs, lists, separators, graphics, tables, table items, etc.”. Thus, rather than teaching applicants’ selecting a plurality of summary entries from the electronic document based on differences in the **content properties**, Reiley teaches only selecting the plurality of summary entries from the electronic document based on the predetermined **content structure**. The “logical structures or elements” of Reiley, namely “headings, paragraphs, lists, separators, graphics, tables, table items, etc.”, are **not** “content properties” as defined in applicants’ claims.

To emphasize this difference, independent claims 1, 10 and 19 have been further amended to define content properties as the Markush group “selected from the group consisting of text formatting, paragraph alignments or indents, and paragraph sizes” (*Ex parte Markush*, 1925 C.D. 126 (Comm'r Pat. 1925)).

Examiner's *Response to Arguments* replies to applicants' argument "A" that Reiley does not disclose "analyzing content properties comprises identifying differences in at least one of text formatting, paragraph alignments or indents, and paragraph sizes" by noting that the claim recitation "at least one of ..." admits of an interpretation of "identifying differences in at least one of text formatting, paragraph alignments or indents, and paragraph sizes" to mean "identifying differences in text formatting, or paragraph alignments or indents, or paragraph sizes or any combination of these". Examiner notes that Reiley discloses wherein said analyzing content properties comprises identifying at least text formatting (col. 7, paragraphs [0077] – [0078]: **identifying nodes** (content elements) that [*sic*] **associated with** bolded text (text formatting) or **nodes associated with** text of a particular length (paragraph size), and identifying a group of **nodes that associate with** patterns of text (text formatting)" (emphasis added). The passage quoted by the Examiner merely emphasizes applicants' argument that Reiley is limited to analyzing documents with a hierarchical structure (i.e. nodes), whereas applicants' claims are directed to methods, program products, servers etc. for "generating summary information" by analyzing documents that have a hierarchical structure (such as the node structure in Reiley), **or** analyzing content properties/formatting where there is no identifiable content structure.. In

Reiley, ‘summary information’ is generated by parsing the Web document to identify a hierarchical organization of elements (i.e. nodes).

For the foregoing reason’s applicants maintain their argument that Reiley does not disclose any analyzing of content properties whatsoever. Reiley teaches only “identifying nodes” (col. 7, paragraph [0077], second line) or “patterns of nodes” (col. 7, paragraph [0078], second line) by analyzing headers or characteristics of the nodes.

Examiner concedes that Reiley et al fails to disclose “wherein said analyzing content properties comprises identifying differences in paragraph sizes”, but cites Bourbakis for disclosing “a method for reconstructing new document from a group of old ones by analyzing paragraphs of the group to extract important statistical feature such as the size of the paragraphs in text characters (Abstract and page 2, paragraph [0024]).” Examiner concludes: “It would have been obvious to one of ordinary skill in the art at the time the invention was made to combine the teachings of Bourbakis and Reiley to include identifying differences in paragraph sizes in the electronic document to remove redundant text”.

First, applicants do not understand how removing redundant text is relevant whatsoever to applicants’ claims. Applicants do not claim removing redundant text. Rather applicants claim “analyzing content properties” by “identifying differences in at least one of text formatting, paragraph alignments or indents, and paragraph sizes”.

Moreover, with great respect, analyzing paragraphs extract statistical feature such as the size of the paragraphs in text characters does not suggest applicants' claimed "generating document summary information" by "selecting the plurality of summary entries from the electronic document based on differences in the content properties". In Bourbakis, content properties of paragraphs are extracted from **multiple documents** and compared for similarity. Similar paragraphs are deemed to be redundant and are removed when preparing the new documents. In applicants' claims, "selecting the plurality of summary entries **from the electronic document**" is "based on differences in the content properties" (i.e. within that electronic document). There are no multiple documents in applicants' claims – only a single document with one of either a predetermined content structure or content properties selected from the group consisting of text formatting, paragraph alignments or indents, and paragraph sizes.

Therefore, applicants' respectfully submit that it would **not** be obvious to combine the teachings of Bourbakis and Reiley to include identifying differences in paragraph sizes **in the electronic document** to remove redundant text".

Examiner's *Response to Arguments* replies to applicants' argument "B" that Reiley does not disclose "selecting the plurality of summary entries from the electronic document based on differences in the content properties" by noting that Reiley discloses on page 3, paragraphs [0038] – [0041]: the web page (electronic document) is divided into several elements including headings, paragraphs, lists, separators, graphics, tables, table item, etc... and these are content properties. The examiner notes that the content transformer "uses analysis rule to categorize the

elements...” Examiner notes further that Reiley discloses retrieving or selecting previously saved transformed content, in which case transformation is not necessary (page 5, paragraph [0056].

First, applicants do not understand how retrieving or selecting previously saved transformed content is relevant whatsoever to applicants’ claims. Applicants do not claim saving and retrieving transformed content. Rather applicants claim “selecting the plurality of summary entries from the electronic document based on differences in the content properties”.

Moreover, as discussed above, examiner has misquoted the passage from Reiley at page 3, paragraphs [0038] – [0041], which in fact reads: “the Web page 205 is divided into several logical structures or elements, including headings, paragraphs, lists, separators, graphics, tables, table items, etc.” Thus, rather than teaching applicants’ selecting a plurality of summary entries from the electronic document based on differences in the **content properties**, Reiley teaches only selecting the plurality of summary entries from the electronic document based on the predetermined **content structure**. The “logical structures or elements” of Reiley, namely “headings, paragraphs, lists, separators, graphics, tables, table items, etc.”, are **not** “content properties” as defined in applicants’ claims.

As discussed above, independent claims 1, 10 and 19 have been further amended to define content properties as the Markush group “selected from the group consisting of text formatting, paragraph alignments or indents, and paragraph sizes” (*Ex parte Markush*, 1925 C.D. 126 (Comm’r Pat. 1925), thereby emphasizing this distinction of Reiley et al

With respect to independent claim 28, the examiner maintains that Reiley discloses a formatted document summarization process which generates summary information by selecting a plurality of summary entries from an electronic document based on differences in content formatting identified in the electronic document. The office action cites paragraphs [0038]-[0041] of Reiley as support for this contention. The office action summarizes this cited passage thusly.

“the web page (electronic document) is divided into several elements including headings, paragraphs, lists, separators, graphics, tables, table item, etc...and these are content properties, and the transformer uses analysis rules to categorize the elements”.

As discussed above in connection with claims 1, 10 and 19, examiner has misquoted the passage from Reiley, which in fact reads: “the Web page 205 is divided into several **logical structures** or elements, including headings, paragraphs, lists, separators, graphics, tables, table items, etc.” Thus, rather than teaching applicants’ selecting a plurality of summary entries from the electronic document based on differences in the **content formatting**, Reiley teaches only selecting the plurality of summary entries from the electronic document based on the predetermined **content structure**. The “logical structures or elements” of Reiley, namely “headings, paragraphs, lists, separators, graphics, tables, table items, etc.”, are **not** “content formatting” as defined in applicants’ claim 28.

To emphasize this difference, independent claim 28 has been further amended to define content formatting as the Markush group “selected from the group consisting of text formatting, paragraph alignments or indents, and paragraph sizes” (*Ex parte Markush*, 1925 C.D. 126 (Comm’r Pat. 1925)).

With respect to independent claim 35, the language of the claim clearly elucidates that the system claimed in the instant application may rely either on the structure of the content of the electronic document or on the formatting of the content of the electronic document in order to generate summary information for electronic documents. As with the previously discussed claims, the subject matter of this claim is not disclosed in the Reiley reference, which operates only on the structure of an electronic document, not on the formatting of the content of an electronic document.

As discussed above in connection with claim 28, examiner has misquoted paragraphs [0038]-[0041] of Reiley, which in fact reads: “the Web page 205 is divided into several logical structures or elements, including headings, paragraphs, lists, separators, graphics, tables, table items, etc.” Thus, rather than teaching applicants’ selecting a plurality of summary entries from the electronic document based on differences in the **content formatting**, Reiley teaches only selecting the plurality of summary entries from the electronic document based on the predetermined **content structure**.

The “logical structures or elements” of Reiley, namely “headings, paragraphs, lists, separators, graphics, tables, table items, etc.”, are **not** “content formatting” as defined in applicants’ claim 35.

To emphasize this difference, independent claim 35 has been further amended to define content formatting as the Markush group “selected from the group consisting of text formatting,

paragraph alignments or indents, and paragraph sizes” (*Ex parte Markush*, 1925 C.D. 126 (Comm’r Pat. 1925)).

For at least these reasons, the cited references do not render unpatentable the subject matter of claims 1, 9, 10, 19, 20, 28, 35 and 36 in the instant application.


New Dependent Claims

New dependent claims 40-42 have been added, having full support in at least paragraphs [0037] – [0039] of the specification as originally filed. According to claims 40-42, the selection of the plurality of summary entries from the electronic document based on differences in the content properties is recited as comprising “determining if the content formatting includes at least one of text formatting, paragraph alignments or indents and in response selecting the plurality of summary entries based on differences in said at least one of text formatting, paragraph alignments or indents; and otherwise selecting the plurality of summary entries based on differences in said paragraph sizes”. This feature permits the selection of summary entries from unstructured formatted electronic documents (i.e. documents containing text formatting, paragraph alignments or indents) or unstructured unformatted electronic documents (i.e. documents that contain no text formatting information, in which case paragraph sizing may be used to select the summary entries).

For at least the reasons set forth above, applicants respectfully submit that claims 1, 9, 10, 19, 20, 28, 35, 36 and 40-42 are patentable over the cited references and are in condition for allowance. Allowance is respectfully requested.

Respectfully Submitted,

JONES DAY



Joseph M. Sauer (Reg. No. 47,919)
Jones Day
NorthPoint, 901 Lakeside Avenue
Cleveland, Ohio 44114
(216)-586-7506